



science + computing ag

Stand und aktuelle Entwicklungsrichtungen im Grid
Computing aus der Perspektive des Technical
Computings in der Industrie

Vortrag von Dr. Peter Stoll

science + computing ag

IT-Dienstleistungen und Software für anspruchsvolle Rechnernetze

Tübingen | München | Berlin | Düsseldorf

Vision: Metapher 'Stromnetz'



Bilder: www.pca.state.mn.us
www.xrfcorp.com
www.dagbladet.no
www.universitas.no

Definition?

1969: Len Kleinrock:

“We will probably see the spread of ‘computer utilities’, which, like present electric and telephone utilities, will service individual homes and offices across the country.”



1998, Carl Kesselman, Ian Foster in

“The Grid: Blueprint for a New Computing Infrastructure.”

“A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities.”

2008 Was ist da? Was kommt?

Aus der Perspektive von science + computing

- Was ist da für das 'High End Computing'?
- Defizite, wo wird gearbeitet?

science + computing ag

- IT-Dienstleistungen und Software für die effiziente Nutzung von Rechnerumgebungen in Forschung und Entwicklung
- Markt des "Technical Computing" (CAD, CAE, CAT)
- über 50 Linux-Cluster in der deutschen Automobil- und anderen Industrien konzipiert
- Partner Platform Computing (LSF), Reseller

Referenzen

Automobilindustrie
 Mikroelektronik
 Luft- und Raumfahrt
 Lifescience
 Öffentlicher Dienst
 IT-Dienstleister



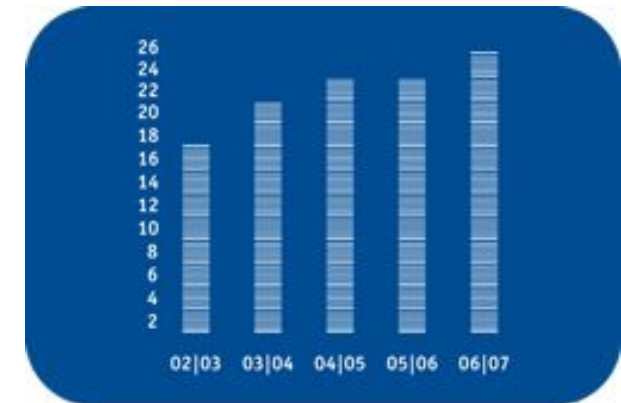
science + computing auf einen Blick

Gründungsjahr 1989

Umsatz 06/07 24,42 Mio €

Mitarbeiter 250

Standorte
Tübingen
München
Berlin
Düsseldorf



Wachstum der Gesamtleistung in Mio. Euro

Partner

Japan: Daikin Industries
Italien: NICE srl
USA: Exa Corporation
Kanada: Platform Computing

1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004



Grid Projekte bei s+c: KnowARC

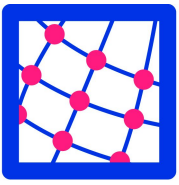


KnowARC

- EU – Forschungsprojekt (STREP)
- Entwicklung von Grid-Middleware
- Software 'ARC'
- 2006 - 2009

Oslo University	Norway
Lund University	Sweden
Copenhagen University	Denmark
Uppsala University	Sweden
NIIF	Hungary
Geneva Uni. Hospitals	Switzerland
Kosice University	Slovakia
Lübeck University	Germany
science+computing ag	Germany
SUN Microsystems	Hungary





Grid-Technologie für kooperative Produktentwicklung am Beispiel der Simulation und des Produktdatenmanagements



Pro**GRID**

gefördert vom

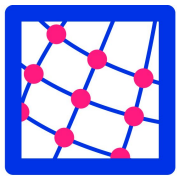
Projekt im D-Grid



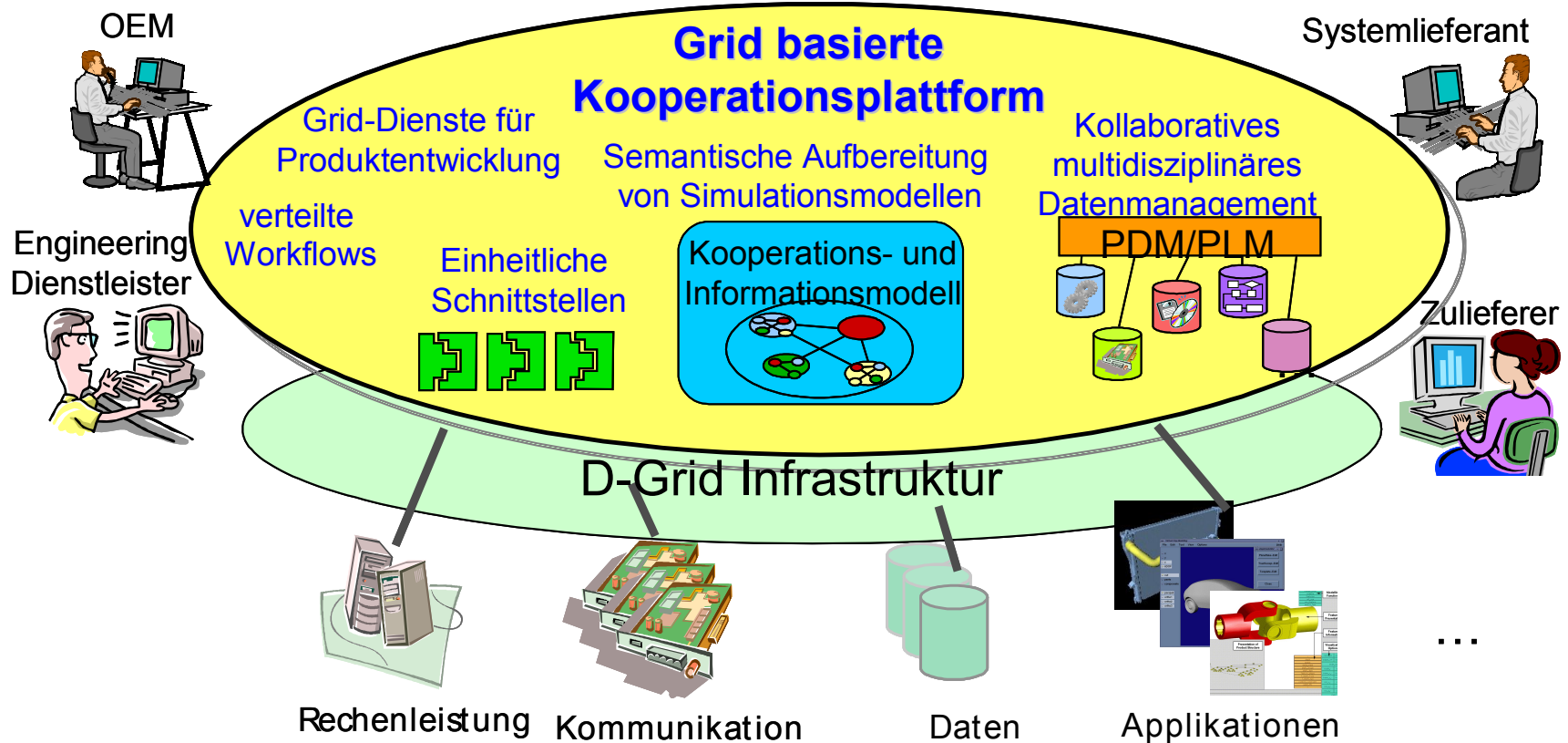
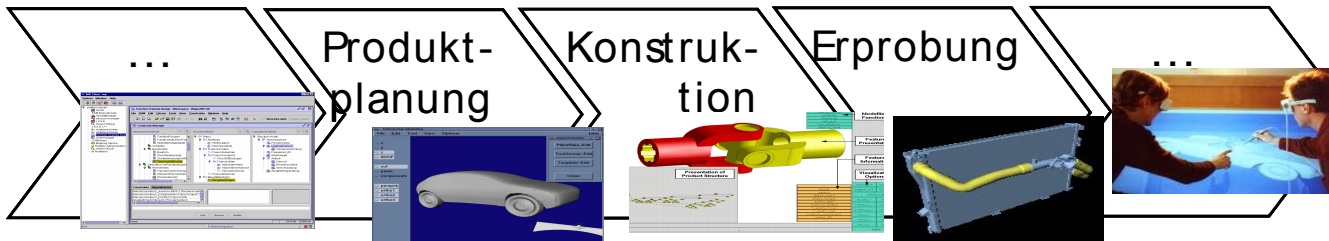
Bundesministerium
für Bildung
und Forschung

Partner

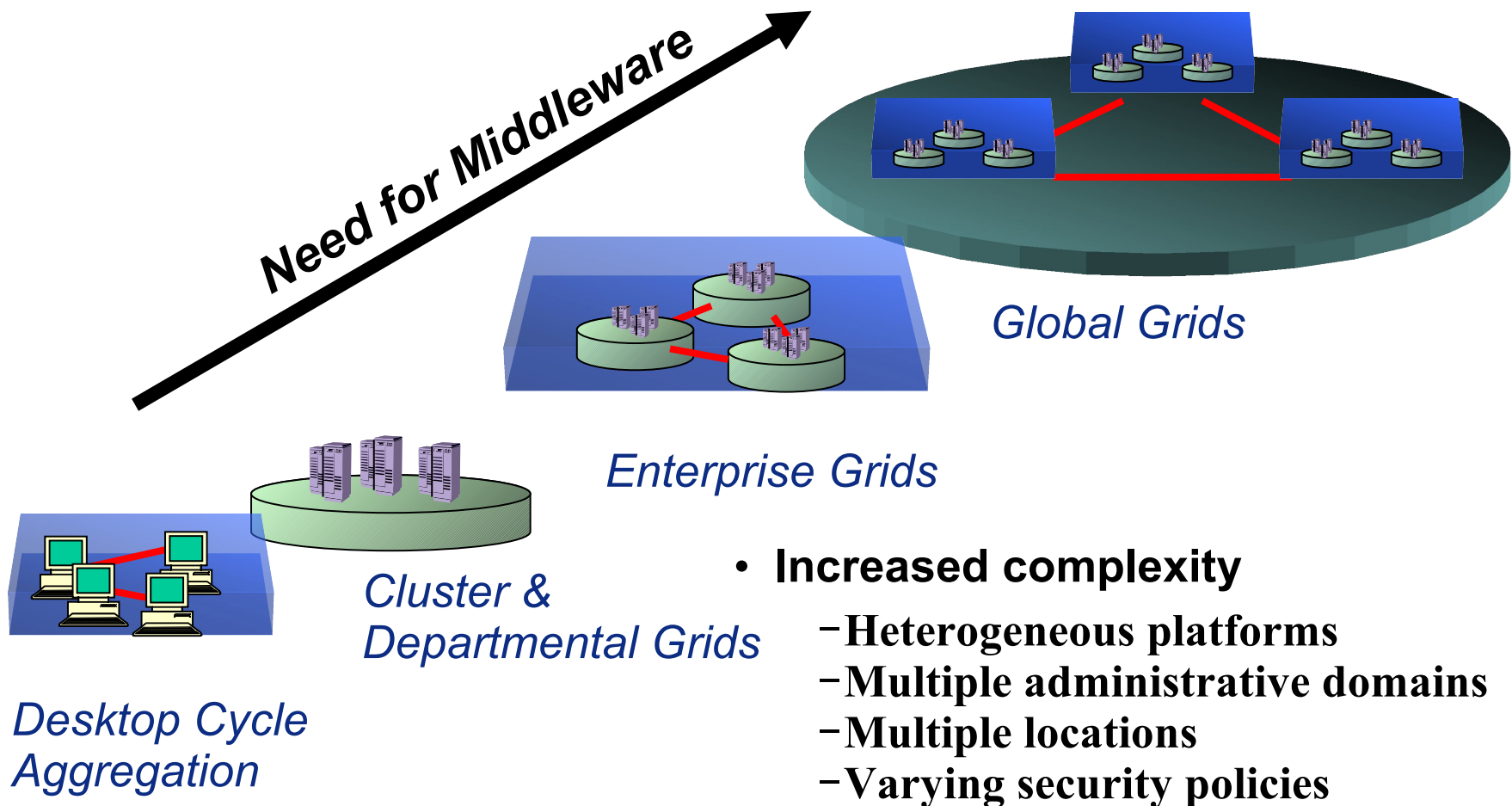
- science + computing ag, Tübingen (Konsortialführer)
- Fraunhofer-Institut für Produktionsanlagen und Konstruktionstechnik (IPK), Berlin
- Fraunhofer-Institut für Algorithmen und Wissenschaftliches Rechnen (SCAI), Sankt Augustin
- FE-Design GmbH, Karlsruhe
- INTES GmbH, Stuttgart
- Wilhelm Karmann GmbH, Osnabrück
- PDTec GmbH, Karlsruhe
- Behr GmbH & Co. KG (assoz)
- DaimlerChrysler AG (assoz)
- Visenso GmbH (assoz)

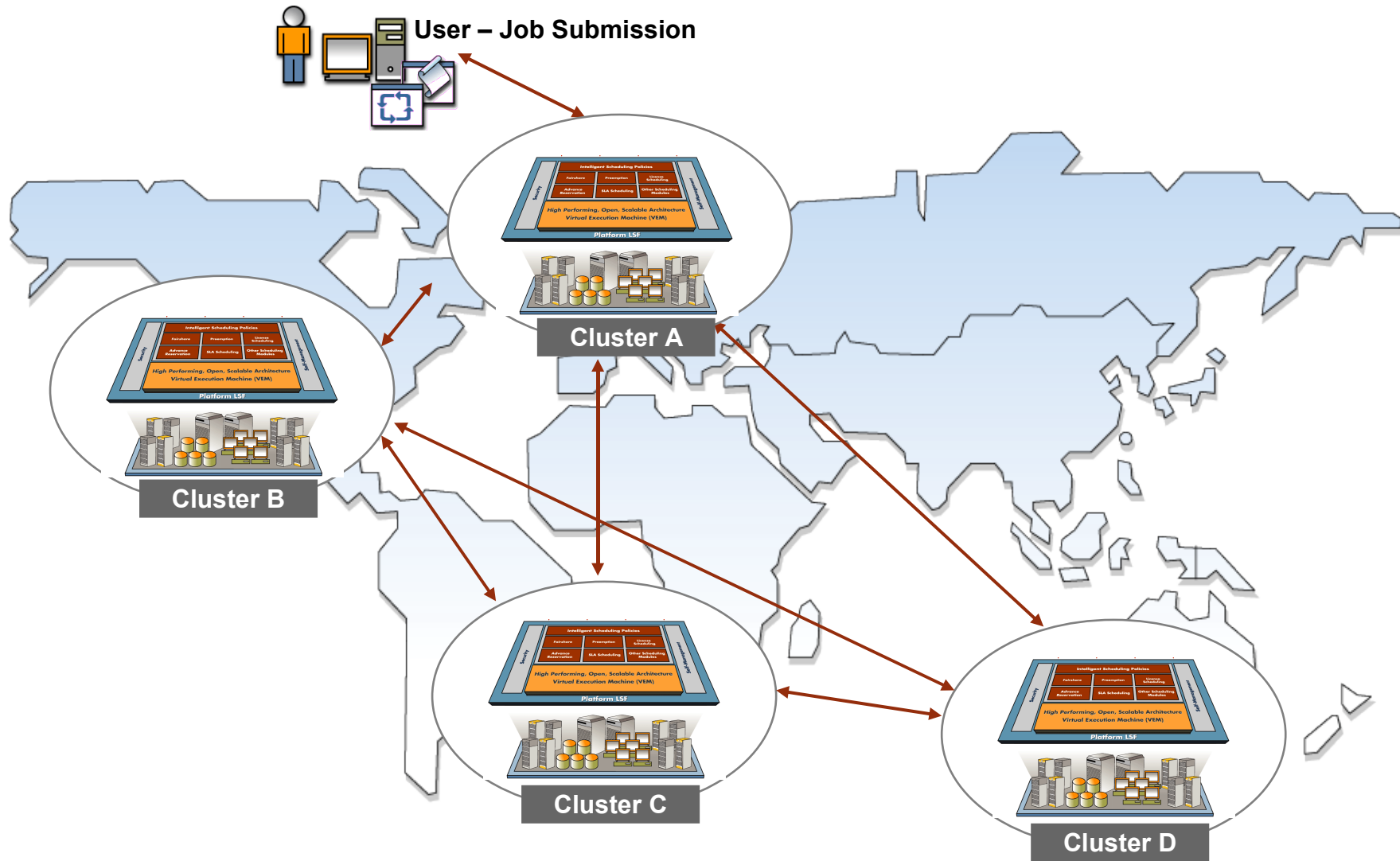


Simulationsdienste in Produktentstehungsprozessen



IT's all about Distributed Resource Management





Case: Infineon Technologies

Anforderung des Kunden

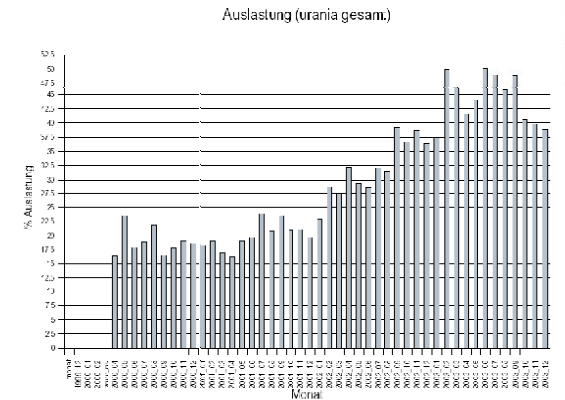
- Grid-Computing im weltweiten Infineon-Netz
- Thin-Client/Compute-Farm-Konzept
- Kostensenkung in Nutzung und Administration

Mehrwert für den Kunden

- **über 100% Steigerung der Rechenauslastung**
- **Steigerung der Arbeitseffizienz**
- **Einsparungen in zweistelliger Millionenhöhe**

Karl Pomschar, CIO Infineon (Handelsblatt, 05.01.2004):

„Bei Infineon wurden bereits Einsparungen in zweistelliger Millionenhöhe realisiert. Ohne Grid wäre Infineon nicht in der Lage, trotz des eingeschränkten IT-Budgets die benötigte Rechenleistung bereitzustellen.“

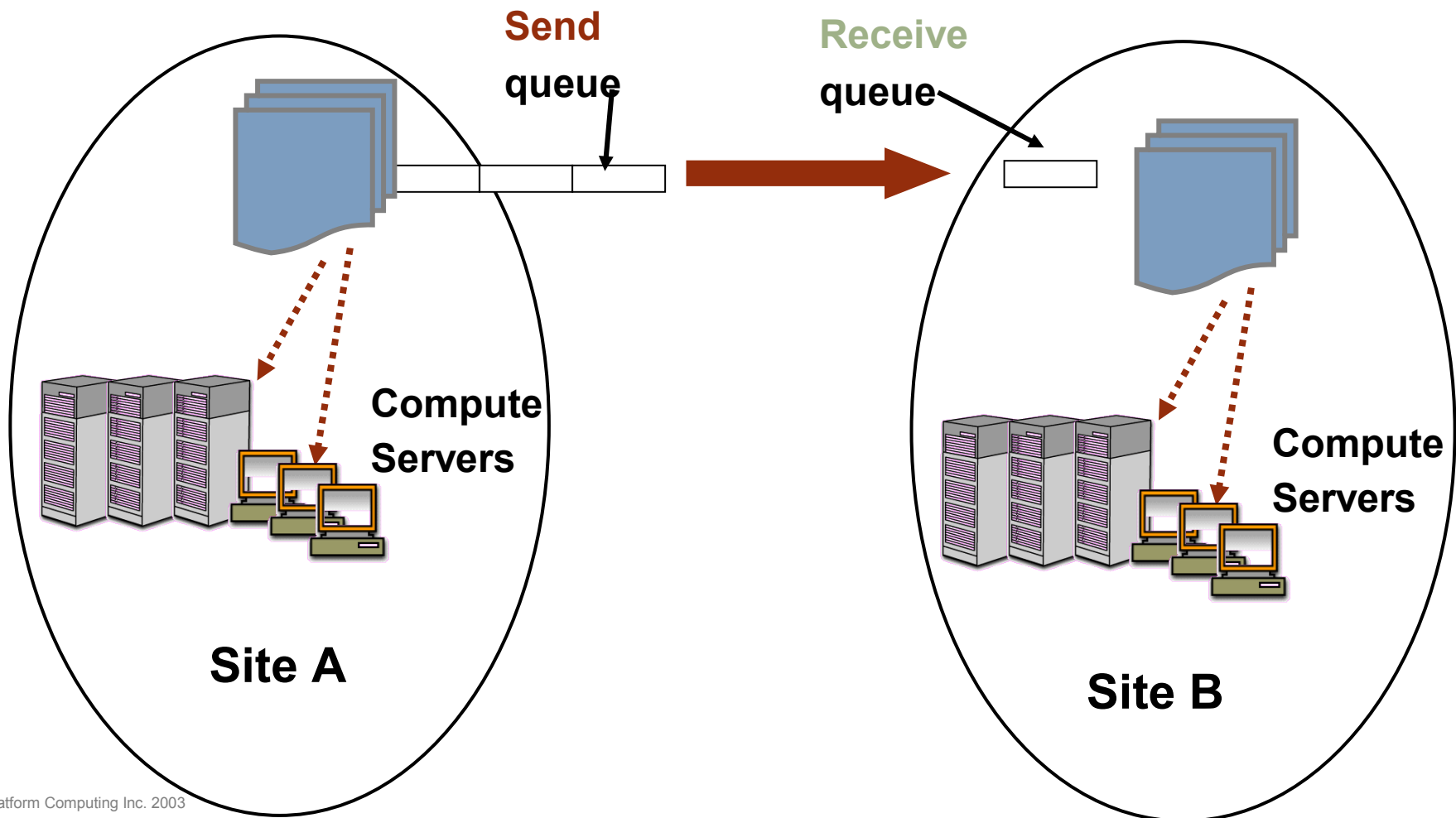


s+c Leistungen

- Distributed Resource Management
- Automatisierung
- HPC



The job forwarding provides the ability to spill jobs over to another cluster as the local cluster becomes busy.



Weltweites Grid?

Das geht heute mit etablierter, kommerzieller Software (z.B. LSF)

- Cluster
- Enterprise Grid
- Verteilung über mehrere Standorte (LSF Multicluster)

Und weiter, das 'weltweite' Grid ?

Was bietet die Grid-Middleware aus den e-Science Bereich?

Klassische (e-Science) Grid Middleware

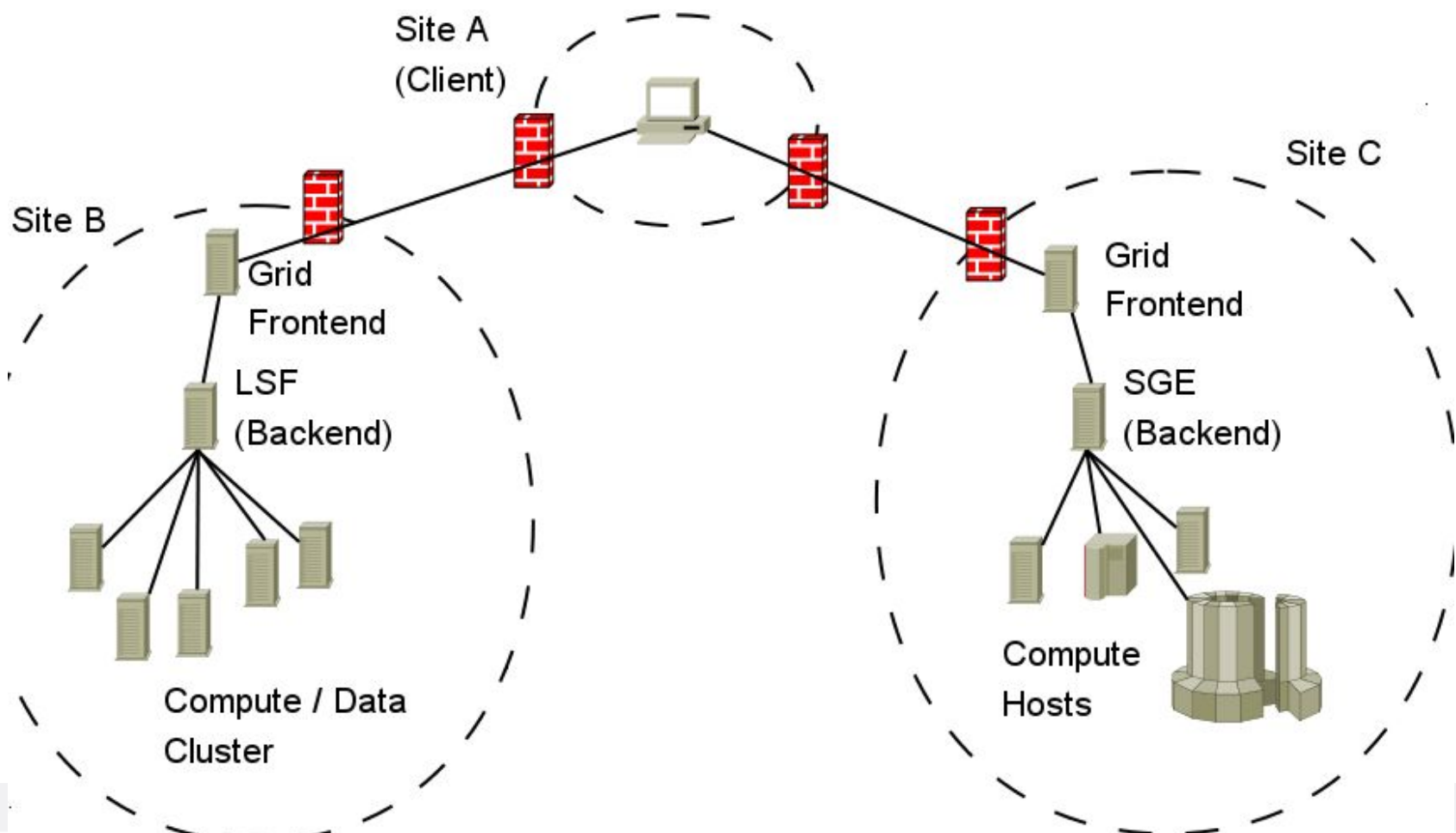
Der Grid Mainstream (nach Fördersumme: xx0 000 000 €)

- Globus
- gLite
- Unicore
- XtremOS
- GRIA
- ARC
- OMII
- ... mehr ...

Entwicklung hauptsächlich akademisch durch Forschungsprojekte
Motiviert durch HEP (Hochenergie-Physik) und ähnliche Probleme

Konzept klassische Grid - Middleware

- Verteilung von Jobs und Daten vom Client über die Grid-Middleware (z.B. Globus) an die Grid-Frontends
- Submit von dort an 'Local Resource Management Systems' (LRMS).



- Prinzip: die Grid-Middleware (Globus ...) fügt eine weitere Ebene über die klassischen Batch-Queing Systeme (LSF ...) ein.
- Unternehmensübergreifende Verteilung von Ressourcen möglich
- Typische Komponenten dieser Systeme:
 - Security
 - Batch-Job-Submit (CPU Cycles)
 - Storage (Datentransfer)
 - Monitoring, Resource discovery
- ***(noch) keine Interoperabilität auf Basis allg. etablierter Standards, bislang nur Gateway - Lösungen***

Hilft das im 'technical computing' in der Industrie?

Im Prinzip ja, aber

Unterschiede in Anforderungen zwischen e-Science und Industrie

	e-Science	Industrie
Benutzer- verwaltung	federated	zentral
Security	Schutz der Ressourcen wichtig, Schutz Daten weniger	Schutz der Daten kritisch
Firewall	weniger streng	strenge policies
Sicht auf Resource	Batch Job / CPU, RAM und Disk	Service / SOA (Einheit von Software und Rechenresource)
Trust	Vertrauen zum Serviceanbieter	Nur begrenztes Vertrauen
Software	Oft selbstkompiliert auf Resource, oder lizenzfrei	Kommerziell, lizenzgebunden

Standards für die wichtigsten Komponenten eines Grid-Systems

- Security

 - X.509 Public Key Infrastructure (PKI) mit Proxy Certificate Profile (RFC3820)

- Batch-Job-Submit (CPU Cycles)

 - OGSA-BES Web-Service mit JSDL (Job Submission Description Language)

- Storage (Datentransfer)

 - GridFTP Protocol

- Monitoring, Resource Discovery

 - weniger weit, GLUE für Resource Discovery

X.509 PKI

- Etabliert: X.509 Public Key Infrastructure (PKI) mit Proxy Certificate Profile (RFC3820)
- Interoperabilität zwischen verschiedenen Grid-Systemen
- Basierend auf GSI (Globus 'Grid Security Infrastructure')
- Public/Private Key – Paar in \$HOME
- Proxy Zertifikat für Single Sign On und Delegation (in /tmp/) das an den Server geschickt wird
- Proxy Zertifikat hat 'Restriktion', derzeit aber nur die Zeit als Restriktion verwendet, obwohl Standard auch mehr ermöglichen würde.

Einschränkung für Betrieb: Volles Vertrauen zum Grid Provider notwendig, da dieser innerhalb der Proxy-Lebensdauer dieses missbrauchen kann.

Grid Standards (2): Job Submit

OGSA-BES + JSDL

- Als Standard verabschiedet:
OGSA-BES (Basic Execution Service) mit JSDL
- BES ist ein Web-Service (WSDL, XSD, ...), Zugriff sprachunabhängig und plattformübergreifend möglich
- Alle wichtigen Grid-Middlewares arbeiten an Implementierungen, aber noch keine fertigen, praxistauglichen Implementierungen
- Job-Beschreibung mit JSDL (Job Submission Description Language), XML-Dialekt. Bereits ohne BES für einige Systeme in Verwendung (z.B. LSF, ARC, ...).

Potentiell sehr hoffnungsvoll, aber noch nicht da.

Risiken: Standard deckt Anforderungen nicht vollständig ab, inkompatible Erweiterungen.

GridFTP

- GridFTP ist etabliert und De-Facto-Standard innerhalb e-Science
- OGSA-Empfehlung, ein RFC Draft existiert
- FTP-Erweiterung für parallelen Datentransfer
- 3rd-party copy
- Verwendet sehr viele Ports um Begrenzungen der TCP-Buffer-Größe bei manchen Systemen zu umgehen (diese begrenzt die Transferrate bei hoher Bandbreite und hoher Latenz).
- D-Grid empfiehlt Ports 20000-25000 für GridFTP offen zu halten, also 5000 Ports!

Technisch interessant, aber wegen Firewall-Problematik häufig nicht einsetzbar.

Grid Standards (4): Monitoring, Resource Discovery

- Für Resource Discovery Aktivitäten in Standardisierungs-Arbeitsgruppe 'GLUE', aber noch ein Stück Weg zu gehen
- Job-Monitoring (Job-Status) teils über BES abgedeckt

- Grid Software für Enterprise Grid existiert kommerziell wie Open Source und ist, teils mit Einschränkungen, einsetzbar.
- Grid Software für Unternehmensübergreifende Grids aus e-Science Bereich prinzipiell geeignet um solche Systeme als Einzellösungen aufzubauen.
- Machbarkeit muss aber im Einzelfall geklärt werden.
- Noch keine allgemein verwendbaren Standards, wie z.B. bei für WWW, mail, dadurch ist jede Lösung eine Einzellösung.

- Trend zu Standardisierung und Interoperabilität setzt sich fort, getrieben durch Ziele der Forschungsförderung
- Konsolidierung der Zahl der weiterentwickelten Grid-Middlewares
- Gemächliche aber stetige Aufnahme industrieller Anforderungen

Stand und aktuelle Entwicklungsrichtungen im Grid Computing aus der Perspektive des Technical Computings in der Industrie



Vielen Dank für Ihre Aufmerksamkeit.

Dr. Peter Stoll

science + computing ag

www.science-computing.de

Telefon 07071 9457-0

info@science-computing.de